

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
2 June 2005 (02.06.2005)

PCT

(10) International Publication Number
WO 2005/050482 A1

(51) International Patent Classification⁷: **G06F 17/30**,
17/00, 12/00

(74) Agent: HANLEY, Mark, G.; GROSSMAN & FLIGHT,
LLC, 20 N. Wacker Drive, Suite 4220, Chicago, IL 60606
(US).

(21) International Application Number:
PCT/US2003/033392

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE,
GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,
MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT,
RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 21 October 2003 (21.10.2003)

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,
SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(25) Filing Language: English

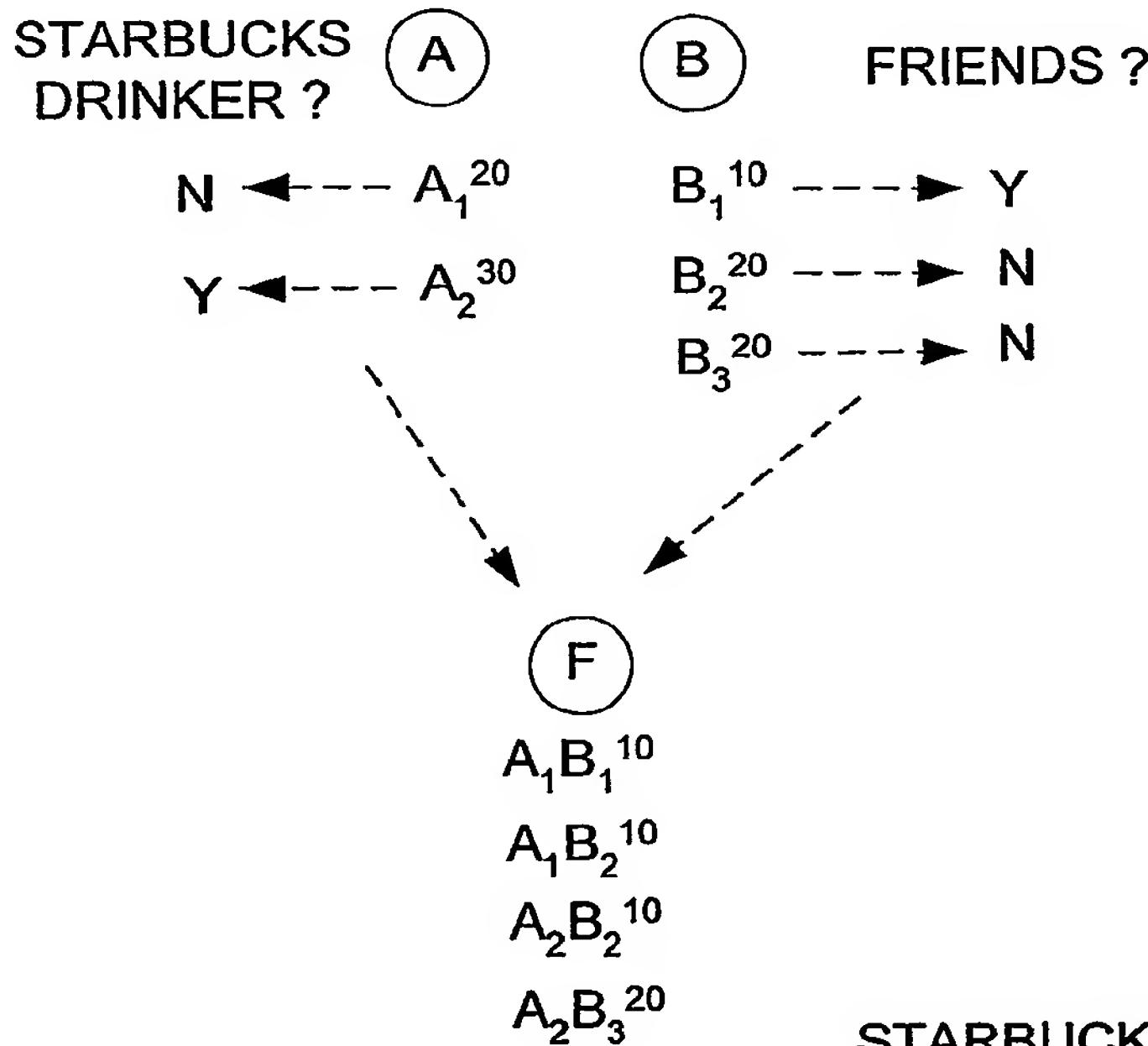
Published:

(26) Publication Language: English

— *with international search report*

[Continued on next page]

(54) Title: METHODS AND APPARATUS FOR FUSING DATABASES



WO 2005/050482 A1

(57) Abstract: Methods and apparatus for fusing multiple databases into a single database are disclosed. A disclosed method determines a ranking of a plurality of matching variables associated with first and second datasets and generates a hierarchical matching grid including a plurality of levels based on the ranking of the plurality of matching variables. The example method identifies first and second sets of match candidates from the first and second datasets based on successive levels of the hierarchical matching grid and fuses records in the first and second sets of match candidates based on probabilities associated with the records.



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.